

OPTIMAL BROWNIAN STOPPING WHEN THE SOURCE AND TARGET ARE RADially SYMMETRIC DISTRIBUTIONS

BY NASSIF GHOUSSEUB[†] YOUNG-HEON KIM[†] AND TONGSEOK LIM[‡]

The University of British Columbia[†] and ShanghaiTech University[‡]

Given two probability measures μ, ν on \mathbb{R}^d , in subharmonic order, we describe optimal stopping times τ that maximize/minimize the cost functional $\mathbb{E}|B_0 - B_\tau|^\alpha$, $\alpha > 0$, where $(B_t)_t$ is Brownian motion with initial law μ and with final distribution –once stopped at τ – equal to ν . Under the assumption of radial symmetry on μ and ν , we show that in dimension $d \geq 3$ and $\alpha \neq 2$, there exists a unique optimal solution given by a non-randomized stopping time characterized as the hitting time to a suitably symmetric *barrier*. We also relate this problem to the optimal transportation problem for *subharmonic martingales*, and establish a duality result. This paper is an expanded version of a previously posted but not published work by the authors [22].

1. Introduction. Let μ and ν be two probability measures on \mathbb{R}^d , $d \geq 2$ with finite first moment, and let $(B_t)_t$ denote the Brownian motion with initial law μ . We consider the following –possibly empty– set of stopping times, with respect to the Brownian filtration:

$$\mathcal{T}(\mu, \nu) = \{\tau \mid \tau \text{ is a stopping time, } B_0 \sim \mu, B_\tau \sim \nu, \text{ and } \mathbb{E}[\tau] < \infty\},$$

where here and in the sequel, the notation $X \sim \lambda$ means that the law of the random variable X is the probability measure λ .

For a given cost function $c : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$, and assuming $\mathcal{T}(\mu, \nu)$ is non-empty, we shall consider the following optimization problem

$$(1.1) \quad \text{Maximize / Minimize } \mathbb{E}[c(B_0, B_\tau)] \quad \text{over } \tau \in \mathcal{T}(\mu, \nu).$$

*The first two authors are partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). Y.-H. Kim was also supported by an Alfred P. Sloan Research Fellowship. T. Lim was partly supported by a doctoral graduate fellowship from the University of British Columbia, by the Austrian Science Foundation (FWF) through grant Y782, and by the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 335421. In addition, T. Lim gratefully acknowledges support from ShanghaiTech University. Part of this research was done while the authors were visiting the Fields institute in Toronto during the thematic program on “Calculus of Variations” in Fall 2014. We are thankful for the hospitality and the great research environment that the institute provided.

MSC 2010 subject classifications: Primary 49-XX, 60-XX; secondary 52-XX

Keywords and phrases: Subharmonic Martingale Optimal Transport, Skorokhod Embedding, Monotonicity, Radial Symmetry.

Recall that a *stopping time* on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$ is a random variable $\tau : \Omega \rightarrow [0, +\infty]$ such that $\{\tau \leq t\} \in \mathcal{F}_t$ for every $t \geq 0$. A *randomized stopping time* is a probability measure τ on $\Omega \times [0, +\infty]$ such that for each $u \in \mathbb{R}_+$, the random time $\rho_u(\omega) := \inf\{t \geq 0 : \tau_\omega([0, t]) \geq u\}$ is a stopping time, where $(\tau_\omega)_\omega$ is a disintegration of τ along the path ω according to \mathbb{P} , that is $\tau(d\omega, dt) = \tau_\omega(dt)\mathbb{P}(d\omega)$. In the sequel, “stopping times” including those in $\mathcal{T}(\mu, \nu)$, will mean possibly randomized stopping times unless stated otherwise. We note that a stopping time is *non-randomized* if the disintegration τ_ω is the Dirac measure on \mathbb{R}_+ for \mathbb{P} a.e. ω . For further details See Section 2, and also e.g. Beiglöck-Cox-Huesmann [2], Guo-Tan-Touzi [26].

The purpose of the present article is to identify and characterize those optimal stopping times for which (1.1) is attained. In particular, we will be investigating when optimal stopping times are unique, ‘true’ as opposed to randomized, and whether they are hitting times of some barriers. Moreover, we consider how conditions of radial symmetry on the source and target measures are reflected in terms of symmetry in the geometry of the barrier.

But first, we recall that the Skorokhod Embedding Problem (SEP) -initiated by Skorokhod [40] in the early 1960s- gives necessary and sufficient condition on a pair (μ, ν) , which insures that the set $\mathcal{T}(\mu, \nu)$ is non-empty. In dimension one, this condition states that μ and ν should be in *convex order*, that is

$$(1.2) \quad \int f d\mu \leq \int f d\nu \text{ for every convex function } f \text{ on } \mathbb{R}^d.$$

Since then, the problem and its variants were investigated by a large number of researchers, and have led to several important results in probability theory and stochastic processes. We refer to Oblój [36] for an excellent survey of the subject, which describes no less than 21 different solutions to (SEP).

More recently, Hobson [27] made the connection between SEP and problems of robust pricing and hedging of financial instruments. Hobson-Klimmek [29] and Hobson-Neuberger [30] connected SEP to finding robust price bounds for the forward starting straddle. See the excellent survey of Hobson [28].

The interest eventually shifted to the problem of finding optimal solutions among those that solve (SEP). In other words, which among the stopping times in $\mathcal{T}(\mu, \nu)$ maximize or minimize a given cost function. Among the multitude of contributions to these questions, we point out the papers of Beiglöck-Cox-Huesmann [2], Cox-Oblój-Touzi [11], Dolinsky-Soner [13, 14], Guo-Tan-Touzi [26], Källblad-Tan-Touzi [32], to name just a few. We do single out, however, the recent work of Beiglöck-Cox-Huesmann [2], which uses the analogy between the optimal SEP and the theory of optimal mass transport, to identify and prove the so-called *Monotonicity Principle* (MP), which will be one of the main tools used in this paper.

Most of the previously mentioned works focus on the one-dimensional case. More recently, Ghoussoub-Kim-Palmer use dynamic programming and PDE methods to

give solutions, that is, optimal non-randomized stopping times, in all dimensions for the minimization case when the cost is either $c(x, y) = |x - y|$ or when it is dynamic and given by a certain Lagrangian [23, 24]. Left open is the case of a general cost, including those of the form

$$(1.3) \quad c(x, y) = |x - y|^\alpha,$$

where $0 < \alpha \neq 2$, or more generally cost functions of the form $c(x, y) = f(|x - y|)$, where $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ is a continuous function such that $f(0) = 0$.

The present paper also deals with higher dimensional situation, but with a focus on the case where both source and target measures are radially symmetric. One of the main goals is to analyze how this symmetry impacts the structure of the barrier sets, whose hitting times determine the optimal stopping times. To state our first main result, we shall denote by $\mathbb{P}_c(O)$ the set of compactly supported probability measures whose supports are in a given open set O of \mathbb{R}^d .

THEOREM 1.1. *Consider the cost function $c(x, y) = |x - y|^\alpha$, where $0 < \alpha \neq 2$, and let $\mu, \nu \in \mathbb{P}_c(\mathbb{R}^d)$ be radial¹, $d \geq 3$ are such that $\mathcal{T}(\mu, \nu)$ is nonempty. Assume $\mu \in H^{-1}(\mathbb{R}^d)$, $\mu \wedge \nu = 0$, $\mu(\{0\}) = 0$, and $\mu(\partial \text{supp } \mu) = 0$.*

Then, there exists a unique solution τ for the minimization/maximization problem (1.1), that is given by a non-randomized stopping time.

Moreover, in the minimization problem with $0 < \alpha \leq 1$, the solution τ is still unique also among all possibly randomized solutions, and without the assumption that $\mu \wedge \nu = 0$

Actually, we shall show that the optimal τ is a hitting time of a suitable barrier. Moreover, each barrier set for the Brownian motion B^x starting from x is axially symmetric around the line that connects the origin and x (see Remark 2.12). A similar structure is detected in the two dimensional case, however, it was not sufficient to yield a non-randomized stopping time (see Remark 2.13).

For the rest of this introduction, we do not assume that μ, ν are necessarily radial, but we shall need the notion of measures in subharmonic order. Recall that a function f is subharmonic on an open set U in \mathbb{R}^d if it is upper-semicontinuous with values on $\mathbb{R} \cup \{-\infty\}$ and satisfies

$$f(x) \leq \frac{1}{\text{Leb}(B)} \int_B f(y) dy \quad \text{for every closed ball } B \text{ in } U \text{ with center at } x.$$

When f is C^2 , the latter condition is equivalent to $\Delta f \geq 0$ on U , where $\Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$ is the Laplacian. We note that, unlike convex functions, a subharmonic function on U may not have any subharmonic extension on a larger domain V . Therefore it is

¹A measure μ is radial if $\mu(A) = \mu(M(A))$ for any $A \in \mathcal{B}(\mathbb{R}^d)$ and any orthogonal matrix M . An example is the Gaussian measure with mean 0 and covariance matrix $\sigma^2 \text{Id}$, $\sigma > 0$.

important to specify a subharmonic function with its domain. We shall therefore consider $\mathcal{SH}(O)$ to be the cone of all subharmonic functions on an open set O .

DEFINITION 1.2. *We say that $\mu, \nu \in \mathbb{P}_c(\mathbb{R}^d)$ are in subharmonic order if the following holds:*

$$(1.4) \quad \int f d\mu \leq \int f d\nu \quad \text{for every } f \in \mathcal{SH}(O),$$

where O is an open set containing $\text{supp}(\mu + \nu)$. We shall then write $\mu \prec_s \nu$.

The fact that in higher dimension, the condition $\mu \prec_s \nu$ is also sufficient for the non-emptiness of $\mathcal{T}(\mu, \nu)$ has been the subject of many studies. See for example Rost [39], Monroe [35], Baxter-Chacon [1], Chacon-Walsh [9], Falkner [17] and many others. We shall give yet another proof of this fact from a different viewpoint, and present the intimate connection between *the optimal Skorokhod embedding problem* (OSEP) and *the subharmonic martingale optimal transport* (SMOT) problem that we now introduce.

For this, recall the *martingale optimal transport* (MOT) problem. Let $\Pi(\mu, \nu)$ be the set of transport plans with *marginals* μ, ν , i.e. the subset of $\mathbb{P}(\mathbb{R}^d \times \mathbb{R}^d)$ consisting of all couplings of μ, ν (see [42]). Then the MOT problem consists of the following:

$$(1.5) \quad \text{Minimize } \text{Cost}[\pi] = \iint c(x, y) d\pi(x, y) \quad \text{over } \pi \in \text{MT}(\mu, \nu),$$

where $\text{MT}(\mu, \nu)$ is the set of *martingale transport plans*, that is, the subset of $\Pi(\mu, \nu)$ such that for each $\pi \in \text{MT}(\mu, \nu)$, its disintegration $(\pi_x)_{x \in \mathbb{R}^d}$ w.r.t. μ satisfies

$$(1.6) \quad f(x) \leq \int f(y) d\pi_x(y)$$

for any convex function f on \mathbb{R}^d . In other words, x is the *barycenter* of π_x , μ -a.e. x .

Problem (1.5) has been extensively studied, especially in one dimension by Beiglöck-Juillet [5], Beiglöck-Nutz-Touzi [6] and more recently in higher dimension by Ghoussoub-Kim-Lim [22]. For more recent developments, see De March-Touzi [12] and Oblój-Siorpaes [37]. In our situation, we need to consider the class of *subharmonic martingale transport plans on O* , that is the class $\text{SMT}_O(\mu, \nu)$ of those $\pi \in \text{MT}(\mu, \nu)$ such that $\text{supp}(\mu + \nu) \subset O$ and the inequality (1.6) also holds for every subharmonic function f on O . This leads us to the *subharmonic martingale optimal transport problem* (on the domain O):

$$(1.7) \quad \text{Minimize } \text{Cost}[\pi] = \iint_{O \times O} c(x, y) d\pi(x, y) \quad \text{over } \pi \in \text{SMT}_O(\mu, \nu).$$

Since every convex function is subharmonic, $\text{SMT}_O(\mu, \nu) \subset \text{MT}(\mu, \nu)$, and for $d = 1$ these two sets are the same. However, for $d \geq 2$, the inclusion is strict and problems (1.5) and (1.7) are different. The following standard example illustrates this difference, and also the importance of the choice of the domain O in the SMOT problem.

EXAMPLE 1.3. Let B_r be the open disk centered at 0 with radius r in \mathbb{R}^2 , and let $S_r = \partial B_r$ be its boundary. Let ρ_r be the uniform probability measure on S_r . As a first example, let $\mu = \rho_1$ and $\nu = \frac{1}{2}(\delta_0 + \rho_2)$. Then clearly $\text{MT}(\mu, \nu) \neq \emptyset$. On the other hand, if we consider the subharmonic function $g(z) = \log |z|$, then

$$\int g(z) d\mu(z) = 0 > -\infty = \int g(z) d\nu(z),$$

which implies that $\text{SMT}_{\mathbb{R}^d}(\mu, \nu) = \emptyset$, since otherwise for any $\pi \in \text{SMT}_{\mathbb{R}^d}(\mu, \nu)$,

$$\int g(y) d\nu(y) = \int \left(\int g(y) d\pi_x(y) \right) d\mu(x) \geq \int g(x) d\mu(x).$$

Another illustrative example consists of taking $\mu = \rho_2$ and $\nu = \rho_3$. Letting $U = \{x \mid 1 < |x| < 4\}$ be an annulus and $V = B_4$, then $\text{SMT}_V(\mu, \nu) \neq \emptyset$ but $\text{SMT}_U(\mu, \nu) = \emptyset$. One way to see this is by considering a Brownian motion B_t with $B_0 \sim \mu$, and letting τ_{B_3} be the first exit time of B_t from B_3 . Then for any $f \in \mathcal{SH}(V)$, $f(B_{t \wedge \tau_{B_3}})$ is a submartingale, and hence the joint law of $(B_0, B_{\tau_{B_3}})$ is an element of $\text{SMT}_V(\mu, \nu)$. On the other hand, the function $h(z) = -\log |z|$ is superharmonic on V and harmonic on U and satisfies $\int h d\mu > \int h d\nu$, which means that $\text{SMT}_U(\mu, \nu) = \emptyset$. In terms of why a stopping time in $\mathcal{T}(\mu, \nu)$ is lacking, it is because since $B_0 \sim \mu$, there is a positive probability that B_t hits the boundary ∂U before hitting S_3 .

The above example shows that, unlike the MOT problem, the SMOT problem is domain sensitive, which leads to the following natural question. Given $\mu, \nu \in \mathbb{P}_c(\mathbb{R}^d)$, for which domain O , does $\text{SMT}_{\mathbb{R}^d}(\mu, \nu) \neq \emptyset$ also imply that $\text{SMT}_O(\mu, \nu) \neq \emptyset$? To provide an answer, we make the following definition.

DEFINITION 1.4. We say that an open set O is a regular domain for $\mu, \nu \in \mathbb{P}_c(\mathbb{R}^d)$ if there exists a compact set K in O such that $\text{supp}(\mu + \nu) \subset K$ and $\mathbb{R}^d \setminus K$ is connected. We denote by $\mathcal{R}(\mu, \nu)$ the set of all regular domains for μ, ν .

Now we define the set of subharmonic martingale transport plans as follows:

$$\text{SMT}(\mu, \nu) = \bigcap_{O \in \mathcal{R}(\mu, \nu)} \text{SMT}_O(\mu, \nu).$$

Also, define Ψ to be a space of continuous test functions for subharmonic martingales

$$\Psi := \{p \in C(\mathbb{R}^d \times \mathbb{R}^d) \mid p(x, x) = 0 \text{ and } p(x, \cdot) \in \mathcal{SH}(\mathbb{R}^d) \text{ for all } x \in \mathbb{R}^d\}.$$

The following theorem tells us that $\text{SMT}(\mu, \nu)$ is not smaller than $\text{SMT}_{\mathbb{R}^d}(\mu, \nu)$, and gives a precise connection between the order of μ, ν , the solvability of SMOT, and that of OSEP.

THEOREM 1.5. *Let $\mu, \nu \in \mathbb{P}_c(\mathbb{R}^d)$. Then*

$$(1.8) \quad \text{SMT}(\mu, \nu) = \text{SMT}_{\mathbb{R}^d}(\mu, \nu) = \left\{ \pi \in \Pi(\mu, \nu) \mid \int p(x, y) d\pi(x, y) \geq 0 \quad \forall p \in \Psi \right\}$$

and the following are equivalent:

- 1) $\mu \prec_s \nu$,
- 2) $\mu \prec_{s(O)} \nu$ for every $O \in \mathcal{R}(\mu, \nu)$,
- 3) $\text{SMT}(\mu, \nu) \neq \emptyset$,
- 4) $\mathcal{T}_O(\mu, \nu) := \mathcal{T}(\mu, \nu) \cap \{\tau \mid \tau = \tau \wedge \tau_O\} \neq \emptyset$ for every $O \in \mathcal{R}(\mu, \nu)$,

where τ_O is the first exit time of a d -dimensional Brownian motion $(B_t)_t$ from O .

Moreover, for each $\pi \in \text{SMT}(\mu, \nu)$ and $O \in \mathcal{R}(\mu, \nu)$, there exists a stopping time τ such that $B_0 \sim \mu$, $B_\tau \sim \nu$, $\tau = \tau \wedge \tau_O$, and π is the joint distribution of (B_0, B_τ) .

We then explore a dual problem by considering the class of functions

$$\mathcal{K}_c = \{(\alpha, \beta, p) \in C_b \times C_b \times \Psi \mid \beta(y) - \alpha(x) + p(x, y) \leq c(x, y)\},$$

and prove the following duality result.

THEOREM 1.6. *Let $\mu, \nu \in \mathbb{P}_c(\mathbb{R}^d)$, and let c be a lower-semicontinuous cost. Then the following duality holds:*

$$\begin{aligned} \sup \left\{ \int \beta d\nu - \int \alpha d\mu \mid (\alpha, \beta, p) \in \mathcal{K}_c \right\} &= \inf \left\{ \int c(x, y) d\pi \mid \pi \in \text{SMT}(\mu, \nu) \right\} \\ &= \inf \{ \mathbb{E}[c(B_0, B_\tau)] \mid \tau \in \mathcal{T}(\mu, \nu) \}. \end{aligned}$$

We remark that a dual attainment result has been achieved recently in [23] for a certain class of cost functions, including the power costs $c(x, y) = \pm|x - y|^\alpha$, $\alpha > 0$, under suitable assumptions as considered in this paper, but without assuming radial symmetry on μ and ν . We in fact use this latter result in the appendix to prove the radial monotonicity principle in Proposition 2.7, which is crucial for this paper.

In section 2 we prove Theorem 1.1 by using that *radial monotonicity principle*. In section 3 we prove Theorems 1.5, and 1.6. Sections 2 and 3 can be read independently.

2. Structure of optimal stopping times in symmetric Skorohod Embeddings. In this section we focus on Theorem 1.1, and for that we start by introducing a symmetric version of the monotonicity principle of Beiglböck-Cox-Huesmann [2], adapted for the radially symmetric case in multi-dimensions.

We first make more precise the filtered Brownian probability space on which we operate. We consider $C(\mathbb{R}_+) = \{\omega : \mathbb{R}_+ \rightarrow \mathbb{R}^d \mid \omega(0) = 0, \omega \text{ is continuous}\}$ to be the path space starting at 0. The probability space will be $\Omega := C(\mathbb{R}_+) \times \mathbb{R}^d$ equipped with the probability measure $\mathbb{P} := \mathbb{W} \otimes \mu$, where \mathbb{W} is the Wiener measure and μ is a given (initial) probability measure on \mathbb{R}^d . Stopping times and randomized stopping times will be with respect to this obviously filtered probability space.

Following Beiglböck-Cox-Huesmann [2], let S be the set of all *stopped paths*

$$(2.1) \quad S = \{(f^x, s) \mid f^x : [0, s] \rightarrow \mathbb{R}^d \text{ is continuous and } f^x(0) = x\},$$

which, by letting $f^x(t) = f^x(s)$ for $t \geq s$, can be viewed as a subset of Ω .

Notice that a (randomized) stopping time ξ can be considered as a probability measure on $\Omega \times [0, +\infty]$. In particular, if the stopping time is finite almost surely, then ξ is concentrated on the set S of stopped paths. In addition, ξ can be disintegrated along the paths in Ω and give a family of probability measures $\{\xi_{\omega^x}\}$ on $[0, +\infty]$, where $\omega^x(t) := \omega(t) + x$.

Now we can also interpret ξ as a transport plan in the following way: for $(f^x, s) \in S$, ξ transports the infinitesimal mass $d\xi((f^x, s))$ from $x \in \mathbb{R}^d$ to $f^x(s) \in \mathbb{R}^d$ along the path (f^x, s) . Now define a map $T : S \rightarrow \mathbb{R}^d \times \mathbb{R}^d$ by

$$(2.2) \quad T((f^x, s)) = (x, f^x(s))$$

and let $T\xi$ be the push-forward of the measure ξ by the map T , thus $T\xi$ is a probability measure on $\mathbb{R}^d \times \mathbb{R}^d$, which is a transport plan.

Next, we introduce the *conditional randomized stopping time given $(f^x, s) \in S$* , that is, the normalized stopping measure given that we followed the path f^x up to time s . For $(f^x, s) \in S$ and $\omega \in C(\mathbb{R}_+)$, define the *concatenated path* $f^x \oplus \omega$ by

$$(f^x \oplus \omega)(t) = \begin{cases} f^x(t) & \text{if } t \leq s, \\ f^x(s) + \omega(t - s) & \text{if } t > s. \end{cases}$$

DEFINITION 2.1 (Conditional stopping time [2]). *Let ξ be a randomized stopping time, defined on Ω . The conditional randomized stopping time of ξ given $(f^x, s) \in S$, denoted by $\xi^{(f^x, s)}$, gives a probability measure on each \mathbb{W} -a.e. $\omega \in C(\mathbb{R}_+)$ as follows:*

$$\begin{aligned} \xi_{\omega}^{(f^x, s)}([0, t]) &= \frac{1}{1 - \xi_{f^x \oplus \omega}([0, s])} (\xi_{f^x \oplus \omega}([0, t + s]) - \xi_{f^x \oplus \omega}([0, s])) \\ &\quad \text{if } \xi_{f^x \oplus \omega}([0, s]) < 1, \\ \xi_{\omega}^{(f^x, s)}(\{0\}) &= 1 \quad \text{if } \xi_{f^x \oplus \omega}([0, s]) = 1. \end{aligned}$$

According to [2], this is the normalized stopping measure of the “bush” which follows the “stub” (f^x, s) . Note that $\xi_{f^x \oplus \omega}([0, s])$ does not depend on the bush ω .

2.1. The Beiglböck-Cox-Huesmann monotonicity principle on stopped paths. Consider the set of all stopped paths S , and let $\Gamma \subset S$ be a concentration set of a given stopping time τ , i.e. $\tau(\Gamma) = 1$. Given the “stopped paths” Γ , consider the “going paths” $\Gamma^< := \{(f, t) \mid \exists (g, t') \in \Gamma, t < t', \text{ and } f \equiv g \text{ on } [0, t]\}$. We shall write

$$(2.3) \quad (x \rightarrow \psi_y : x' \rightarrow y') \in (\tau, \Gamma),$$

if there exist $(f^x, t) \in \Gamma^<$, $(g^{x'}, t') \in \Gamma$ such that $y = f^x(t)$, $y' = g^{x'}(t')$, and $\psi_y = \text{Law}(B^y(\tau(f^x, t)))$. Notice that the measure ψ_y is the conditional probability measure generated by the strong Markov property of the Brownian motion that once stopped at y at time t then continued following the stopping time rule τ . Also, the fact $(g^{x'}, t') \in \Gamma$ means intuitively that the path $g^{x'}$ drops a mass at the point y' under the stopping rule τ . We shall consider the following cost for such a pair of transport

$$(2.4) \quad C(x \rightarrow \psi_y : x' \rightarrow y') = \int c(x, z) d\psi_y(z) + c(x', y').$$

In the sequel, we shall denote by $\text{SH}(x)$ the set of probability measures ψ whose barycenter x satisfies $\delta_x \prec_s \psi$. As seen in Theorem 1.5, these are the probabilities that can be obtained by stopped Brownian motions starting at x . The following principle will be crucial for the proof of Theorem 1.1. It was proved by Beiglböck-Cox-Huesmann [2] for more general path-dependent costs.

THEOREM 2.2 (Monotonicity principle [2]). *Suppose c is a cost function and that τ is an optimal stopping time for the minimization problem (1.1). Then, there exists a Borel set $\Gamma \subseteq S$ such that $\tau(\Gamma) = 1$, and Γ is c -monotone in the following sense: If $\psi_y \in \text{SH}(y)$ and $(x \rightarrow \psi_y : x' \rightarrow y) \in (\tau, \Gamma)$, then*

$$(2.5) \quad C(x \rightarrow \psi_y : x' \rightarrow y) \leq C(x \rightarrow y : x' \rightarrow \psi_y).$$

Here is a first easy but informative application of this principle, but first an important result about the case where the source and target have overlapping mass. Recall that stopping measures (which are not necessarily probability measures) lie in $C(\mathbb{R}_+; \mathbb{R}^d) \times [0, +\infty]$.

PROPOSITION 2.3. *Let τ_0 be a zero-stopping measure, i.e. concentrated on $C(\mathbb{R}_+; \mathbb{R}^d) \times \{0\}$, and assume that its projection on the second component \mathbb{R}^d is $\mu \wedge \nu$. Assume the cost function is of the form $c(x, y) = f(|x - y|)$ and that it satisfies*

$$(2.6) \quad c(x, x) = 0 \quad \text{and} \quad c(x, z) \leq c(x, y) + c(y, z) \text{ for all } x, y, z \in \mathbb{R}^d,$$

with equality in the triangle inequality occurring if and only if x, y, z are along a line.

Then for any $\tau \in \mathcal{T}(\mu, \nu)$ that minimizes problem (1.1), we must have $\tau_0 \leq \tau$, and therefore $\tau^* := \tau - \tau_0$ solves the minimization problem (1.1) with the same cost but for the disjoint source marginal $\bar{\mu} := \mu - \mu \wedge \nu$, and target marginal $\bar{\nu} := \nu - \mu \wedge \nu$.

PROOF. Note that having $\tau_0 \leq \tau$ means that under τ , the common mass $\mu \wedge \nu$ stays put. Let $\pi := T\tau$ (resp., $\pi_0 := T\tau_0$) be a coupling of the pair (μ, ν) (resp., $(\mu \wedge \nu, \mu \wedge \nu)$) and observe that

$$\tau_0 \leq \tau \iff \pi_0 \leq \pi.$$

We will prove the latter, and for this we largely follow the argument of [34]. Let $\Gamma \subset S$ be a c -monotone set for τ in Theorem 2.2. Let $T_\Gamma := T(\Gamma)$ so that $\pi(T_\Gamma) = 1$. Let π_D be the restriction of π on the diagonal $\Delta = \{(x, x) : x \in \mathbb{R}^d\}$. Note that $\text{supp } \pi_0 \subset \Delta$. Now if $\pi_0 \not\leq \pi$, the measure $\pi_0 - \pi_D$ has a non-zero positive part. Let η be the push-forward measure of this positive part by the projection $(x, x) \mapsto x$. Denote $d\pi(x, y) = d\pi_x(y)d\mu(x)$ and $d\pi(x, y) = d\pi^y(x)d\nu(y)$ to be disintegrations of π w.r.t. μ and ν respectively. Then by definition of η , we have

$$\pi_x \neq \delta_x \text{ and } \pi^x \neq \delta_x, \quad \eta - a.e. x.$$

This implies that there exist $(f^x, 0) \in \Gamma^<$, $(g^z, t) \in \Gamma$ such that $z \neq x$, $x = g^z(t)$ and $\pi_x = \text{Law}(B_\tau^x)$. Now recall the assumption on c that $c(y, x) + c(x, z) \geq c(y, z)$, and the inequality is strict unless z, x, y lie on a line in this order. Hence

$$\int c(y, x) d\pi_x(y) + c(x, z) > \int c(y, z) d\pi_x(y).$$

As $c(x, x) = 0$, this means

$$C(x \rightarrow \pi_x : z \rightarrow x) > C(x \rightarrow x : z \rightarrow \pi_x),$$

a contradiction to Theorem 2.2. \square

REMARK 2.4. Note that the cost $c(x, y) = |x - y|^\alpha$, $0 < \alpha \leq 1$ satisfies the triangle inequality (2.6), and the above proposition implies that whenever we are studying the minimization problem with such a cost, we can assume that $\mu \wedge \nu = 0$ without loss of generality. Indeed, the proposition says that in this case all the optimal (possibly randomized) stopping times τ of the minimization problem (1.1) are uniquely decomposed into two stopping measures as $\tau = \tau_0 + \tau^*$, where τ_0 is concentrated at the time $T = 0$ while τ^* is the stopping measure at positive times. Furthermore, given marginals μ and ν , τ_0 is the largest possible stopping measure in such a way that τ_0 acts as an identity transport from $\mu \wedge \nu$ to itself. In Section 2.4 we will prove that the measure τ^* , once disintegrated along each path, is supported at a single time, i.e. it is non-randomized. Thus, the optimal stopping times $\tau = \tau_0 + \tau^*$ can be randomized only at time 0.

2.2. The radially symmetric monotonicity principle. We now give a variant of Theorem 2.2, which exploits the radial symmetry of the marginals μ and ν . For this, we will introduce several notions related to radial symmetry. First, we give the definition of R -equivalence. Recall the definition of the transport plan $T\xi$ induced by the stopping time ξ (see (2.2)).

DEFINITION 2.5. *Let $\lambda(x) = |x|$ be the modulus map.*

1. *Two probability measures φ and ψ on \mathbb{R}^d are said to be R -equivalent if their push-forward measures by λ coincide, i.e. $\lambda_{\#}\varphi = \lambda_{\#}\psi$. We then write $\varphi \cong_R \psi$.*
2. *Two stopping times ξ and ζ are said to be R -equivalent if the first marginals and second marginals of $T\xi$ and $T\zeta$ are R -equivalent, respectively. We then write $\xi \cong_R \zeta$.*

The following symmetrization was introduced in [34] for the *Martingale Optimal Transport Problem* with radially symmetric marginals. It will also be useful in this paper. Let \mathfrak{M} be the group of all $d \times d$ real orthogonal matrices, and let \mathcal{H} be the Haar measure² on \mathfrak{M} . For a given $M \in \mathfrak{M}$ and a stopping time ξ , we define $M\xi$ as follows: for each $A \subset S$, set

$$(M\xi)(A) = \xi(M^{-1}(A)).$$

Clearly, $M\xi$ is also a stopping time. Now we introduce the symmetrization operator which acts on both the probability measures on \mathbb{R}^d and on the stopping times.

DEFINITION 2.6. *The symmetrization operator Θ acts on the set of probability measures on \mathbb{R}^d , and on the set of stopping times as follows:*

1. *For each probability measure μ on \mathbb{R}^d and $B \subset \mathbb{R}^d$,*

$$(\Theta\mu)(B) = \int_{M \in \mathfrak{M}} (M\mu)(B) d\mathcal{H}(M).$$

2. *For each stopping time ξ and $A \subset S$,*

$$(\Theta\xi)(A) = \int_{M \in \mathfrak{M}} (M\xi)(A) d\mathcal{H}(M).$$

Observe that $\Theta\mu$ is the unique radially symmetric probability measure which is R -equivalent to μ . Moreover, for any stopping time ξ , notice that (see [34] for a proof)

$$(2.7) \quad \text{if } T\xi \text{ has marginals } \mu \text{ and } \nu, \text{ then } T(\Theta\xi) \text{ has marginals } \Theta\mu \text{ and } \Theta\nu.$$

²The Haar measure \mathcal{H} associated to a compact topological group G is the unique probability measure which is left- and right-invariant: $H(gA) = H(A)$ and $H(Ag) = H(A)$ for every $g \in G$ and $A \in \mathcal{B}(G)$. See [18, Chapter 11] for more details.

This leads to the following important observation: Assume $c(x, y)$ is a rotation invariant cost function, i.e., $c(Mx, My) = c(x, y)$ for any $M \in \mathfrak{M}$, and define the cost of a stopping time ξ to be:

$$C(\xi) = \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y) T\xi(dx, dy).$$

If ξ solves the minimization problem (1.1) where $T\xi$ has radially symmetric marginals μ and ν , then for any stopping time ζ , we must have

$$(2.8) \quad C(\zeta) \geq C(\xi) \quad \text{whenever} \quad \zeta \cong_R \xi.$$

Indeed, if $C(\zeta) < C(\xi)$, then since $C(\zeta) = C(\Theta\zeta)$, $\Theta\zeta$ will solve the minimization problem (1.1) with the same marginals μ, ν and less cost, a contradiction. Of course, if ξ solves the maximization problem then the opposite inequality in (2.8) must hold.

We are now ready to introduce the radial monotonicity principle. A proof will be provided in the appendix.

PROPOSITION 2.7 (Radial monotonicity principle). *Let $c(x, y) = \pm|x - y|^\alpha$, $\alpha > 0$, and $\mu, \nu \in \mathbb{P}_c(\mathbb{R}^d)$ be radially symmetric. Assume that $\mu \in H^{-1}(\mathbb{R}^d)$, $\mu \prec_s \nu$, $\mu \wedge \nu = 0$, and $\mu(\partial \text{supp } \mu) = 0$. Suppose τ is an optimal stopping time for the corresponding minimization problem (1.1). Then, there exists a Borel set $\Gamma \subseteq S$ such that $\tau(\Gamma) = 1$, and Γ is radially c -monotone in the following sense:*

If $\psi_y \in \text{SH}(y)$, $(x \rightarrow \psi_y : x' \rightarrow y') \in (\tau, \Gamma)$ and $|y| = |y'|$, then

$$(2.9) \quad C(x \rightarrow \psi_y : x' \rightarrow y') \leq C(x \rightarrow y : x' \rightarrow \varphi_{y'})$$

for any $\varphi_{y'} \in \text{SH}(y')$ that is R -equivalent to ψ_y .

REMARK 2.8. The result should hold for all radially symmetric measures without the additional assumptions on μ and ν , i.e., $\mu \in H^{-1}(\mathbb{R}^d)$ and $\mu(\partial \text{supp } \mu) = 0$.

Actually, we believe the principle of radial monotonicity should follow directly from the general monotonicity principle (Theorem 2.2) once applied to the case where the marginals are radially symmetric. An idea for the proof goes as follows: If the optimal stopping time τ allows a particle starting at x to diffuse when it reaches y so that it becomes the probability measure ψ_y , but takes another particle at x' to stop at y' , and if we have the opposite inequality

$$(2.10) \quad C(x \rightarrow \psi_y : x' \rightarrow y') > C(x \rightarrow y : x' \rightarrow \varphi_{y'})$$

for some $\varphi_{y'} \in \text{SH}(y')$, then we can modify the stopping time τ to τ' in such a way that the particle at x now stops at y by τ' , but instead the particle at x' starts diffusing at y' until it becomes $\varphi_{y'}$. Then (2.10) means that the cost for τ' is smaller

than that of τ , but note that the modified stopping time τ' may not satisfy the terminal marginal condition ν . However, as $\psi_y \cong_R \varphi_{y'}$ and $|y| = |y'|$, τ' is also R -equivalent to τ , a contradiction by (2.8). The radial monotonicity principle asserts the existence of a set of stopped paths Γ which supports the optimal stopping time and resists any such modification.

2.3. A variational lemma. In this section we prove our key observation, namely, Lemma 2.9, which shows a variational calculus with elements in $\text{SH}(y)$. This will yield Proposition 2.10 that provides crucial comparisons that are used in the next section to prove our main theorem. To explain the idea, first let $S_{y,r}$ be the uniform probability measure on the sphere of center y and radius r , arguably the most simple element in $\text{SH}(y)$. We choose $c(x, y) = |x - y|$ for simplicity and consider the following “gain” function,

$$G(x) = G(x, y, r) := \int |x - z| dS_{y,r}(z) - |x - y|,$$

and its gradient $\nabla_x G$. Note that G is essentially the increase in cost when the Dirac mass at y diffuses uniformly onto the sphere. It satisfies the following properties:

1. If $|x - y| < |x' - y'|$ and r is fixed, then $G(x, y, r) > G(x', y', r)$.
In other words, for the same diffusion, the increase in cost is greater when the distance $|x - y|$ is small. Hence, for the minimization problem, it is better to stop particles near the source x , and let the particles that are far from x to diffuse, as long as the given marginal condition is respected.
2. The vector $\nabla G(x)$ points toward the direction $y - x$, thus the directional derivative $\nabla_u G(x)$ is

$$\nabla_u G(x) < 0 \quad \text{if} \quad \langle u, y - x \rangle < 0.$$

Hence the gain function decreases when x moves away from the “center of diffusion” y . By combining this with the monotonicity principle (Proposition (2.7)), one can get crucial information about the optimal stopping time.

From now on, $c(x, y) = |x - y|^\alpha$, $0 < \alpha \neq 2$. We define the *gain function*

$$G(x, \psi_y) := \int |x - z|^\alpha d\psi_y(z) - |x - y|^\alpha \quad \text{for} \quad \psi_y \in \text{SH}(y).$$

Note that G is essentially the increase in cost when the Dirac mass at y diffuses to ψ_y . The following variational lemma is crucial for our analysis.

LEMMA 2.9. *Let x, y be nonzero vectors in \mathbb{R}^d and let $r = |x|$. Let u be a unit tangent vector to the sphere S_r at x , such that $\langle u, y \rangle < 0$. Let $\psi_y \in \text{SH}(y) \setminus \{\delta_y\}$.*

Then, there exists a $\varphi_y \in \text{SH}(y)$ which is R -equivalent to ψ_y , such that

$$\begin{aligned} G(x, \varphi_y) &= G(x, \psi_y) \\ \nabla_u G(x, \varphi_y) &< 0 \quad \text{if } 0 < \alpha < 2, \\ \nabla_u G(x, \varphi_y) &> 0 \quad \text{if } \alpha > 2, \end{aligned}$$

where the directional derivative is applied to the x variable.

PROOF. Let $c(x, z) = |x - z|^\alpha$ and take its partial derivative $\frac{\partial}{\partial x_d}$ at $x = 0$, to obtain

$$(2.11) \quad h(z) := \nabla_{e_d} c(x, z)|_{x=0} = -\alpha |z|^{\alpha-2} z_d$$

Taking the Laplacian in z , we get

$$(2.12) \quad \Delta h(z) = -\alpha(\alpha - 2)(\alpha + d - 2) |z|^{\alpha-4} z_d.$$

We see that the function h is

- i) strictly superharmonic in the lower half-space $\{z_d < 0\}$ if $0 < \alpha < 2$
- ii) strictly subharmonic in the lower half-space $\{z_d < 0\}$ if $\alpha > 2$.

Let $0 < \alpha < 2$, and assume without loss of generality that $x = x_1 e_1 = (x_1, 0, \dots, 0)$ and $u = e_d$. Then $\langle u, y \rangle < 0$, which means that y is in the lower half-space. Let $H = \{z \in \mathbb{R}^d : z_d = 0\}$ and choose a stopping time τ for the Brownian motion B^y such that $\text{Law}(B_\tau^y) = \psi_y$, and let η be the first time B^y hits H . Let $\sigma_y = \text{Law}(B_{\tau \wedge \eta}^y)$. Then σ_y is supported in the closed lower half-space and it is nontrivial, hence by the strict superharmonicity (2.12), we have

$$\nabla_u G(x, \sigma_y) < 0.$$

Now we modify τ to τ' in the following way; if $\tau \leq \eta$, then we let $\tau' = \tau$. But if $\tau > \eta$, in other words if a particle at y has landed on H but not completely stopped by τ , then we symmetrize the remaining time of τ (i.e. the conditional stopping time) with respect to H and get τ' . More precisely, let τ_H be the reflection of the conditional stopping time of τ with respect to H ; that is, if τ stops a path emanating from H , then τ_H stops the reflected path at the same time. Now define $\tau' := \frac{\tau + \tau_H}{2}$ to be a randomization; before re-starting Brownian motion on H , we flip a coin and choose either τ or τ_H for the conditional stopping time.

Now, define $\varphi_y = \text{Law}(B_{\tau'}^y)$ and observe that

- i) $G(x, \varphi_y) = G(x, \psi_y)$ and $\varphi_y \cong_R \psi_y$, by the symmetry with respect to H in the definition of τ' .
- ii) $\nabla_u G(x, \varphi_y) = \nabla_u G(x, \sigma_y)$, since the function $z \mapsto \nabla_{e_d} c(x, z)$ is odd in z_d (see (2.11)) hence the symmetrization in the definition of τ' does not change $\nabla_u G$.

This proves the lemma. \square

Notice that the above lemma implies, in particular for $0 < \alpha < 2$, that

$$0 < G(x, \psi_y) - G(z, \varphi_y) = C(x \rightarrow \psi_y : z \rightarrow y) - C(x \rightarrow y : z \rightarrow \varphi_y)$$

if $z = x + \epsilon u$ with $\langle u, y \rangle < 0$, for small $\epsilon > 0$, where ϵ depends on x and ψ_y .

In the following proposition, we make this result less infinitesimal. This will be crucial in the next section to prove our main theorem.

PROPOSITION 2.10. *Let $\varphi_y \in \text{SH}(y) \setminus \{\delta_y\}$ be given, and define*

$$\underline{G}(x) := \min_{\sigma_y \in \text{SH}(y), \sigma_y \cong_R \varphi_y} \int |x - z|^\alpha d\sigma_y(z) - |x - y|^\alpha,$$

$$\overline{G}(x) := \max_{\sigma_y \in \text{SH}(y), \sigma_y \cong_R \varphi_y} \int |x - z|^\alpha d\sigma_y(z) - |x - y|^\alpha.$$

Then $\underline{G}, \overline{G}$ are attained, and for $x_0, x_1, y \in \mathbb{R}^d$ with $|x_0| = |x_1|$ and $|x_0 - y| < |x_1 - y|$,

$$\begin{aligned} \underline{G}(x_0) &> \underline{G}(x_1) \quad \text{and} \quad \overline{G}(x_0) > \overline{G}(x_1) \quad \text{if} \quad 0 < \alpha < 2, \\ \underline{G}(x_0) &< \underline{G}(x_1) \quad \text{and} \quad \overline{G}(x_0) < \overline{G}(x_1) \quad \text{if} \quad \alpha > 2. \end{aligned}$$

PROOF. Define

$$\begin{aligned} \mathfrak{R}(x, \varphi_y) &:= \left\{ \psi_y \in \text{SH}(y) : \psi_y \in \arg \min_{\sigma_y \cong_R \varphi_y} \int |x - z|^\alpha d\sigma_y(z) \right\}, \\ \overline{\mathfrak{R}}(x, \varphi_y) &:= \left\{ \psi_y \in \text{SH}(y) : \psi_y \in \arg \max_{\sigma_y \cong_R \varphi_y} \int |x - z|^\alpha d\sigma_y(z) \right\}. \end{aligned}$$

We will soon show that these are nonempty. Note that, by definition

$$\begin{aligned} \underline{G}(x) &= G(x, \psi_y) = \int |x - z|^\alpha d\psi_y(z) - |x - y|^\alpha \quad \text{for any } \psi_y \in \mathfrak{R}(x, \varphi_y), \\ \overline{G}(x) &= G(x, \psi_y) = \int |x - z|^\alpha d\psi_y(z) - |x - y|^\alpha \quad \text{for any } \psi_y \in \overline{\mathfrak{R}}(x, \varphi_y). \end{aligned}$$

First, we claim that $\underline{G}(x)$ and $\overline{G}(x)$ are continuous.

Indeed, let $x_n \rightarrow x$ in \mathbb{R}^d , and define

$$\underline{C}(x) := \underline{G}(x) + |x - y|^\alpha = \int |x - z|^\alpha d\psi_y(z) \quad \text{for any } \psi_y \in \mathfrak{R}(x, \varphi_y).$$

Set $a_n = \underline{C}(x_n)$ and $a = \underline{C}(x)$. Choose any subsequence $\{a_k\}$ of $\{a_n\}$ and a corresponding sequence of measures $\psi_k \in \mathfrak{R}(x_k, \varphi_y)$. By compactness, $\{\psi_k\}$ has a subsequence $\{\psi_l\}$ which converges to, say ψ . Note that $\psi \in \text{SH}(y)$ and $\psi \cong_R \varphi_y$ by weak

convergence. Now, write

$$\begin{aligned} \int |x_l - z|^\alpha d\psi_l(z) - \int |x - z|^\alpha d\psi(z) &= \left[\int (|x_l - z|^\alpha - |x - z|^\alpha) d\psi_l(z) \right] \\ &\quad + \left[\int |x - z|^\alpha d\psi_l(z) - \int |x - z|^\alpha d\psi(z) \right]. \end{aligned}$$

The first bracket goes to zero as $l \rightarrow \infty$ since $|x_l - z|^\alpha - |x - z|^\alpha \rightarrow 0$ uniformly on every compact set in \mathbb{R}^d , and the second bracket goes to zero since $\psi_l \rightarrow \psi$.

Now we claim that

$$(2.13) \quad \int |x - z|^\alpha d\psi(z) = \underline{C}(x) = a, \quad \text{i.e.} \quad \psi \in \underline{\mathfrak{R}}(x, \varphi_y).$$

If not, then there exists a $\rho \in \underline{\mathfrak{R}}(x, \varphi_y)$ such that

$$\begin{aligned} \int |x - z|^\alpha d\psi(z) &> \int |x - z|^\alpha d\rho(z), \quad \text{hence} \\ \int |x_l - z|^\alpha d\psi_l(z) &> \int |x_l - z|^\alpha d\rho(z) \quad \text{for all large } l, \end{aligned}$$

a contradiction since $\psi_l \in \underline{\mathfrak{R}}(x_l, \varphi_y)$. Therefore, $a_n \rightarrow a$, since this holds for any subsequence.

To complete the proof of the proposition, we let $0 < \alpha < 2$. Note that x_0, x_1, y are nonzero. Let $r = |x_0| = |x_1|$. Without loss of generality, we can assume that there exists a differentiable curve $x(t) : [0, 1] \rightarrow S_r$ with $x(0) = x_0, x(1) = x_1$ such that $|x(t) - y|$ is strictly increasing. In other words, $x(t)$ satisfies that $|x(t)| = r$ and $\langle \frac{d}{dt}x(t), y \rangle < 0$ for all t . We note that, although finding such a curve is not always possible, we can always choose an alternative point $x'_1 = Mx_1$ for some orthogonal matrix M with $My = y$, so that there is a geodesic curve on the sphere S_r connecting x_0 and x'_1 with the desired property. Notice that $\underline{G}(x_1) = \underline{G}(x'_1)$, hence such a change does not affect the conclusion.

Now for a fixed $t \in [0, 1]$, choose any $\psi_y \in \underline{\mathfrak{R}}(x(t), \varphi_y)$. Then Lemma 2.9 gives $\sigma_y \in \underline{\mathfrak{R}}(x(t), \varphi_y)$ with $\frac{d}{dt}G(x(t), \sigma_y) < 0$. By definition, $\underline{G}(x(s)) \leq G(x(s), \sigma_y)$ for any s , and $\underline{G}(x(t)) = G(x(t), \sigma_y)$. Hence

$$\begin{aligned} \limsup_{\epsilon \downarrow 0} \frac{\underline{G}(x(t+\epsilon)) - \underline{G}(x(t))}{\epsilon} &\leq \limsup_{\epsilon \downarrow 0} \frac{G(x(t+\epsilon), \sigma_y) - G(x(t), \sigma_y)}{\epsilon} \\ &= \frac{d}{dt}G(x(t), \sigma_y) < 0. \end{aligned}$$

The function $\underline{G}(x(t))$ is continuous and satisfies the above strict inequality for each $t \in [0, 1]$, hence it must be strictly decreasing.

For $\overline{G}(x(t))$, we similarly use $\sigma_y \in \overline{\mathcal{R}}(x(t), \varphi_y)$ and $\frac{d}{dt}G(x(t), \sigma_y) < 0$ to get

$$\begin{aligned} \liminf_{\epsilon \downarrow 0} \frac{\overline{G}(x(t-\epsilon)) - \overline{G}(x(t))}{\epsilon} &\geq \liminf_{\epsilon \downarrow 0} \frac{G(x(t-\epsilon), \sigma_y) - G(x(t), \sigma_y)}{\epsilon} \\ &= -\frac{d}{dt}G(x(t), \sigma_y) > 0. \end{aligned}$$

We again see that $\overline{G}(x(t))$ is strictly decreasing.

The case $\alpha > 2$ can be proved in a similar fashion, and the proposition follows. \square

2.4. Optimal Stopping problem for radially symmetric marginals. Finally, armed with Lemma 2.9 and Proposition 2.10, we establish Theorem 1.1.

Proof of Theorem 1.1. We give a proof for the minimization problem for the case $0 < \alpha < 2$; the case $\alpha > 2$, or the maximization problem, can be proved similarly.

Fix $0 < \alpha < 2$, and let τ be a minimizer for (1.1). For $x, y \neq 0$ in \mathbb{R}^d with $x \nparallel y$, (here \parallel denote parallelism), we define the barrier set:

$$U_x^y = \{z \in \mathbb{R}^d : |z| = |y| \text{ and } \langle x, z \rangle > \langle x, y \rangle\}.$$

The set U_x^y looks like the spherical cap of radius $|y|$, which is symmetric around the axis in the direction of x , containing the point y in its boundary.

We shall say that a pair (f, s) and (g, t) in S is *forbidden* if

$$f(0) = g(0) \neq 0 \text{ and } \exists s' < s \text{ such that } f(s') \in U_{g(0)}^{g(t)}.$$

In words, a forbidden pair consists of a path that penetrates the barrier generated by the other path. We let \mathbf{FP} denote the set of forbidden pairs.

First, we show that there exists $\Gamma \subset S$ on which τ is concentrated, such that Γ does not admit any “forbidden pair” that lies in $\Gamma \times \Gamma$. Indeed, choose the α -monotone Γ for τ as given by Proposition 2.7 and suppose that $\mathbf{FP} \cap (\Gamma \times \Gamma) \neq \emptyset$, i.e. there exists a forbidden pair (f, s) and (g, t) in Γ where (f, s) penetrates the barrier $U_{g(0)}^{g(t)}$, that is, not stopping when it hits the set. Then, by the Markov property, the penetrating path (f, s) yields a nontrivial subharmonic measure, say $\psi_y \in \text{SH}(y)$, namely the conditional probability, whose barycenter is at the point $y = f^x(s')$ where the barrier $U_{g(0)}^{g(t)}$ is hit. But this contradicts Proposition 2.10 and Proposition 2.7. Hence, $\mathbf{FP} \cap (\Gamma \times \Gamma) = \emptyset$.

Now, we show that since Γ does not allow forbidden pairs, then every stopping time concentrated on Γ is necessarily non-randomized, which clearly yields the uniqueness. Indeed, let ξ be any stopping time in $\mathcal{T}(\mu, \nu)$ with $\xi(\Gamma) = 1$. Define $\Gamma_0 = \{(f^x, s) \in \Gamma \mid s = 0\}$, i.e. Γ_0 consists of the paths that are stopped immediately at time 0. Let $\Gamma_+ = \Gamma \setminus \Gamma_0$. We can assume $x \neq 0$ for every $(f^x, s) \in \Gamma$ as $\mu(\{0\}) = 0$. Moreover,

because $d \geq 3$, the probability of the Brownian motion from x to hit a line segment is zero, so we can assume that $f^x(s) \not\parallel x$.³ Let ξ_0 be the restriction of ξ on Γ_0 , and ξ_+ be on Γ_+ .

We claim that since $\mathbf{FP} \cap (\Gamma \times \Gamma) = \emptyset$, ξ_+ must be of non-randomized type.

Let us suppose the contrary. Then there exists an element $(f^x, s) \in \Gamma$, $s > 0, x \neq 0$, such that the conditional stopping time $\xi^{(f^x, s)}$ is nonzero. This means that the Brownian motion which has followed the path f^x up to time $s > 0$ will continue its motion at $y := f^x(s)$. Now consider the barrier $U := U_x^y$ and note that the Brownian motion starting at y governed by any non-zero stopping time will go through the surface U before its complete stop, since y is on the boundary of $(d-1)$ -dimensional set U . This implies that there is a stopped path $(g^y, t) \in S$ such that the concatenation $(f^x \oplus g^y, s+t) \in \Gamma$ and for some $s' < s+t$, $(f^x \oplus g^y)(s') \in U$. This means that the pair $((f^x \oplus g^y, s+t), (f^x, s)) \in \Gamma \times \Gamma$ is a forbidden pair, which is a contradiction.

The separation assumption $\mu \wedge \nu = 0$ implies $\xi_0 = 0$, yielding that ξ is non-randomized. This implies the uniqueness of ξ in the usual way, that is if τ and τ' are two minimizers and if their disintegrations do not agree, i.e. $\tau_{\omega^x} \neq \tau'_{\omega^x}$ for all $\omega^x \in B$ with $\mathbb{P}(B) > 0$, then the stopping time $\frac{\tau + \tau'}{2}$, which is obviously a minimizer, must be of randomized type, thus yielding a contradiction. This concludes the proof. \square

REMARK 2.11. In fact the proof of Theorem 1.1 shows that, if the radial monotonicity principle holds without the assumption $\mu \wedge \nu = 0$, then every optimal stopping time τ is decomposed into two stopping measures as $\tau = \tau_0 + \tau^*$, where τ_0 is supported at time $T = 0$ while τ^* is supported in $T > 0$ and is non-randomized.

REMARK 2.12. Let Γ be the c -monotone set as given in Proposition 2.7 on which the optimizer in Theorem 1.1 is concentrated. The proof of Theorem 1.1 in fact tells us that, for the minimization problem with $0 < \alpha < 2$ or the maximization problem with $\alpha > 2$, the optimal stopping time is given by the first time Brownian motion B^x hits the following union of barriers

$$\mathcal{U}_x := \cup_y U_x^y, \text{ where } y = f^x(s), y \not\parallel x, \text{ for some } (f^x, s) \in \Gamma.$$

Moreover, by the uniqueness property and the radial symmetry of μ and ν , the sets \mathcal{U}_x 's are congruent under rotation, that is if M is an orthogonal matrix and $M(x) = x'$, then $M(\mathcal{U}_x) = \mathcal{U}_{x'}$.

For minimization problem with $\alpha > 2$ or maximization problem with $0 < \alpha < 2$, we have the same type of result, but the barrier will be reversed: it will be given by

$$\mathcal{V}_x := \cup_y V_x^y, \text{ where } y = f^x(s), y \not\parallel x, \text{ for some } (f^x, s) \in \Gamma,$$

³The assumption $d \geq 3$ is used here only.

where V_x^y is the reversed barrier

$$V_x^y = \{z \in \mathbb{R}^d : |z| = |y| \text{ and } \langle x, z \rangle < \langle x, y \rangle\}.$$

This is due to the interchange of the superharmonic and subharmonic region of the derivative of the gain function (2.11), according to the value of α . Also note that when dealing with maximization problem, the inequalities (2.5) and (2.9) in the monotonicity principles must be reversed.

REMARK 2.13. Observe that in the above proof, the argument breaks down in the two dimensional case, because the probability for a Brownian motion to hit a line segment is not zero when the dimension ≤ 2 . Still, it shows that the optimal (possibly randomized) stopping time τ has to stop the Brownian path completely once it stops at a point not parallel to the initial point. Therefore, in particular, for the minimization problem with $0 < \alpha < 2$, the set

$$\mathcal{U}_x := \cup_y U_x^y, \text{ where } y = f^x(s), y \not\parallel x, \text{ for some } (f^x, s) \in \Gamma$$

is a barrier. However, there are chances that the Brownian path may stop only partially at points parallel to x , then to continue until it hits the set \mathcal{U}_x . Similar result holds for the maximization problem or with $\alpha > 2$.

Finally, we note that the ideas in this section can be applied to costs that are more general than the ones of the form $|x-y|^\alpha$ considered in this paper. In particular, they apply to a class of cost functions $c(x, y)$ that are invariant under rotation and whose directional derivatives $\nabla_u c(x, y)$ in the x -variable have suitable sub/superharmonic regions in the y -variable.

3. Subharmonic martingale optimal transport problem. In the previous section we focused on the structure of optimal stopping times, utilizing the radial symmetry of marginals. From now on we do not assume the radial symmetry but consider general marginals. We assume compact support, but see Proposition 3.4 for an exception.

As mentioned in the introduction, a subharmonic function on an open subset of \mathbb{R}^d may not allow a subharmonic extension on all of \mathbb{R}^d in general. When can $f \in \mathcal{SH}(O)$ be approximated by $\tilde{f} \in \mathcal{SH}(\mathbb{R}^d)$ in O ? To give an answer, we recall the following.

LEMMA 3.1 (Walsh [43]). *Let K be a compact subset of \mathbb{R}^d such that $\mathbb{R}^d \setminus K$ is connected. Then for each $\varepsilon > 0$ and a harmonic function u on an open set containing K , there exists a harmonic polynomial v such that $|u - v| < \varepsilon$ on K .*

Keeping this in mind, we turn to the proof of Theorem 1.5. First we need the following very likely known lemma. Given an open set O in \mathbb{R}^d , let $C(O)$ be the

space of continuous functions on O . We give a topology on $C(O)$ which is induced by the following convergence, namely, $f_n \rightarrow f$ iff $\|f_n - f\|_{L^\infty(K)} \rightarrow 0$ on every compact subset K of O . It is clear that this topology is metrizable via the following metric

$$d(f, g) := \sum_{n=1}^{\infty} 2^{-n} \min(1, \|f - g\|_{K_n})$$

where $\{K_n\}$ is a compact exhaustion of O ; K_n is compact, $K_n \subset \text{int}(K_{n+1})$, and $\cup_n K_n = O$.

LEMMA 3.2. *The space $\mathcal{H}(O)$ of harmonic functions on O , is separable under d .*

PROOF. The space $C(O)$ is separable under the metric d , e.g. the set $\mathbb{Q}[x_1, \dots, x_d]$ of polynomials with rational coefficients is a countable dense subset of $C(O)$ by Stone-Weierstrass Theorem. Then since every subspace of a separable metric space is separable, $\mathcal{H}(O)$ is separable as well. \square

PROOF OF THEOREM 1.5. To prove (1.8)), note that if $\pi \in \text{SMT}_{\mathbb{R}^d}(\mu, \nu)$, then for any $p \in \Psi$ we have

$$\int p(x, y) d\pi(x, y) = \int p(x, y) d\pi_x(y) d\mu(x) \geq \int p(x, x) d\mu(x) = 0$$

in view of the subharmonicity of $y \mapsto p(x, y)$. Hence we have the following inclusions

$$\text{SMT}(\mu, \nu) \subset \text{SMT}_{\mathbb{R}^d}(\mu, \nu) \subset \left\{ \pi \in \Pi(\mu, \nu) \mid \int p(x, y) d\pi(x, y) \geq 0 \quad \forall p \in \Psi \right\}.$$

Now let $\pi \in \Pi(\mu, \nu)$ and assume $\int p d\pi \geq 0$ for every $p \in \Psi$. We now prove that $\pi \in \text{SMT}(\mu, \nu)$. Indeed, let $\mathbf{1}_{B_{z,r}}$ be the indicator function on the ball $B_{z,r}$ of center z and radius r , and consider the functions of the form

$$p(x, y) = \mathbf{1}_{B_{z,r}}(x)(\varphi(y) - \varphi(x)) \quad \text{where} \quad \varphi \in \mathcal{H}(\mathbb{R}^d).$$

Then $p(x, y) \in \Psi$ ⁴. For $z \in \text{supp } \mu$, we have

$$\begin{aligned} 0 &\leq \frac{1}{\mu(B_{z,r})} \iint p(x, y) d\pi(x, y) = \frac{1}{\mu(B_{z,r})} \int_{B_{z,r}} \left(\int \varphi(y) d\pi_x(y) - \varphi(x) \right) d\mu(x) \\ &\rightarrow \int \varphi(y) d\pi_z(y) - \varphi(z) \quad \text{as} \quad r \rightarrow 0, \quad \mu - a.e. z, \end{aligned}$$

⁴Technically $\mathbf{1}_{B_{z,r}}(x)$ is not continuous, but it can be approximated by continuous and compactly supported functions.

where the convergence is due to e.g. [33, Lemma 4.1.2.]. By changing $\varphi \mapsto -\varphi$ in $p(x, y)$, we deduce

$$(3.1) \quad \int \varphi(y) d\pi_z(y) = \varphi(z) \quad \mu - a.e. z.$$

One may notice that the μ - a.e. convergence set could depend on the choice of φ , but the application of Lemma 3.2 then ensures that the equality (3.1) holds for all $\varphi \in \mathcal{H}(\mathbb{R}^d)$, μ - a.e..

Let ψ be the fundamental solution of the Laplace equation, and let $\psi_a(x) = \psi(x - a)$ and $\psi_{a,c}(x) = \max(\psi(x - a), c)$ where $a \in \mathbb{R}^d$ and $c \in \mathbb{R}$. Note that $\psi_{a,c}$ is continuous and subharmonic on \mathbb{R}^d . By applying the above argument with $p(x, y) = \mathbf{1}_{B_{z,r}}(x)(\psi_{a,c}(y) - \psi_{a,c}(x))$ and letting $c \searrow -\infty$, we deduce

$$(3.2) \quad \int \psi_a(y) d\pi_z(y) \geq \psi_a(z) \quad \text{for every } a \in \mathbb{R}^d, \mu - a.e. z.$$

Now let O be a regular domain for μ, ν , and let K be a compact set in O such that $\text{supp}(\mu + \nu) \subset K$ and $\mathbb{R}^d \setminus K$ is connected. Let O' be an open precompact subset of O containing K . For each $f \in \mathcal{SH}(O)$, there exists a nonnegative Borel measure κ on O such that f can be decomposed as

$$f(x) = h(x) + \int_{O'} \psi_a(x) d\kappa(a) \quad \forall x \in O', \text{ for some } h \in \mathcal{H}(O')$$

by Riesz representation theorem. Let h_ε be a harmonic polynomial such that $|h - h_\varepsilon| < \varepsilon$ on K by Walsh's theorem. Then for μ -a.e. x , we have

$$\begin{aligned} \int f(y) d\pi_x(y) &= \int h(y) d\pi_x(y) + \int_{\mathbb{R}^d} \int_{O'} \psi_a(y) d\kappa(a) d\pi_x(y) \\ &\geq \int h_\varepsilon(y) d\pi_x(y) + \int_{O'} \int_{\mathbb{R}^d} \psi_a(y) d\pi_x(y) d\kappa(a) - \varepsilon \\ &\geq h_\varepsilon(x) + \int_{O'} \psi_a(x) d\kappa(a) - \varepsilon \\ &\geq h(x) + \int_{O'} \psi_a(x) d\kappa(a) - 2\varepsilon = f(x) - 2\varepsilon. \end{aligned}$$

Letting $\varepsilon \rightarrow 0$ we get $\int f(y) d\pi_x(y) \geq f(x)$ μ -a.e. x , implying $\pi \in \text{SMT}_O(\mu, \nu)$. As $O \in \mathcal{R}(\mu, \nu)$ was arbitrary we deduce that $\pi \in \text{SMT}(\mu, \nu)$. This completes the proof of the equality (1.8).

Now we prove the equivalence in Theorem 1.5. Notice that (1) \iff (2) is immediate by the same approximation argument as above. The direction (3) \implies (1) is also immediate, as for $f \in \mathcal{SH}(\mathbb{R}^d)$ and $\pi \in \text{SMT}(\mu, \nu)$,

$$\int f(y) d\nu(y) = \int f(y) d\pi_x(y) d\mu(x) \geq \int f(x) d\mu(x).$$

To see that (4) \implies (3), let O be an open ball containing $\text{supp}(\mu + \nu)$, and take $\tau \in \mathcal{T}_O(\mu, \nu)$. Then for $f \in \mathcal{SH}(\mathbb{R}^d)$, since $(f(B_{t \wedge \tau}))_t$ is a submartingale, we have

$$\mathbb{E}[f(Y) \mid X] \geq f(X), \quad \text{where } X = B_0 \sim \mu, Y = B_\tau \sim \nu.$$

This means that the joint law of (X, Y) belongs to $\text{SMT}(\mu, \nu)$.

It remains to show the implication (2) \implies (4). This will be done in Proposition 3.4 below, where we prove it for possibly non-compactly supported marginals.

Finally, let now $\pi \in \text{SMT}(\mu, \nu)$ and its disintegration $(\pi_x)_x$ with respect to μ . Let $(B_t)_t$ be a Brownian motion with $B_0 \sim \mu$, and let $O \in \mathcal{R}(\mu, \nu)$ be bounded. Since $\delta_x \prec_{s(O)} \pi_x$ for μ -almost all x , and noting that B_0 and $B_t - B_0$ are independent, one can apply the implication (2) \implies (4) to the subharmonic ordered pair (δ_x, π_x) , and use the measurable selection theorem, to find a stopping time $\tau = \tau \wedge \tau_O$ such that given B_0 , we have $\text{Law}(B_\tau \in \cdot \mid B_0) = \pi_{B_0}(\cdot)$. It is then clear that $\text{Law}(B_0, B_\tau) = \pi$. \square

The following is now immediate.

COROLLARY 3.3. *Let $\mu, \nu \in \mathbb{P}_c(\mathbb{R}^d)$ and assume $\mu \prec_s \nu$. Then*

$$\inf \{ \mathbb{E}[c(B_0, B_\tau)] \mid \tau \in \mathcal{T}(\mu, \nu) \} = \inf \left\{ \iint c(x, y) d\pi \mid \pi \in \text{SMT}(\mu, \nu) \right\}.$$

The following proposition may have its own interest. Let $\mathbb{P}_2(O)$ be the set of probability measures concentrated in O and with finite second moments. Define the order $\mu \prec_{s_2(O)} \nu$ by

$$\mu \prec_{s_2(O)} \nu \iff \int f d\mu \leq \int f d\nu \quad \text{for every } f \in \mathcal{SH}(O) \text{ with } f(x) \leq C_f(1 + |x|^2).$$

PROPOSITION 3.4. *Let $\mu, \nu \in \mathbb{P}_2(O)$ and assume $\mu \prec_{s_2(O)} \nu$. Then $\mathcal{T}_O(\mu, \nu) \neq \emptyset$.*

REMARK 3.5. For the proposition, we need not assume O is regular since the proof clearly implies that if $\mu, \nu \in \mathbb{P}_c(O)$ are such that $\mu \prec_{s(O)} \nu$, then necessarily $\mathcal{T}_O(\mu, \nu) \neq \emptyset$.

To prove the proposition, we first introduce the notion of *spherical martingales*.

DEFINITION 3.6. *Let U be the uniform probability measure on the unit sphere S^{d-1} in \mathbb{R}^d . Let $(X_i)_{i=1}^\infty$ be i.i.d random variables on some probability space (Ω, \mathbb{P}) , whose distribution is U . A stochastic process $(F_n)_{n \geq 0}$ is called a spherical martingale (valued) in O if there is an associated sequence of bounded measurable functions*

$$r_n : \mathbb{R}^d \times \underbrace{S^{d-1} \times \dots \times S^{d-1}}_{(n-1)\text{-times}} \rightarrow \mathbb{R}_+$$

such that

$$F_n(X_1, \dots, X_n) - F_{n-1}(X_1, \dots, X_{n-1}) = r_n(F_0, X_1, \dots, X_{n-1}) \cdot X_n$$

and if for every $n \in \mathbb{N}$, $0 \leq \lambda \leq 1$ and (X_1, X_2, \dots, X_n) ,

$$F_{n-1}(X_1, \dots, X_{n-1}) + \lambda r_n(F_0, X_1, \dots, X_{n-1}) X_n \in O.$$

At each time n , a particle splits uniformly onto a surrounding sphere of radius r_n . An important observation is that the push-forward measure of δ_x by a spherical martingale F_n has the same law as B_τ^x , where the stopping time τ is defined as follows: τ_1 is the first time B^x hits the sphere $S(x, r_1(x))$ centered at x and with radius $r_1(x)$. If $B_{\tau_1}(\omega) = x_1 \in S(x, r_1(x))$, then define τ_2 to be the first hitting time B^{x_1} hits the sphere $S(x_1, r_2(x, x_1))$. One can then define inductively a sequence of stopping times $\tau_1 \leq \tau_2 \leq \dots$ such that $F_n(\mathbb{P}) = \text{Law}(B_{\tau_n}^x)$.

The following lemma is an analogue of a result of Bu-Schachermayer [8, Proposition 2.1] about analytic martingales.

LEMMA 3.7. *Let O be an open set in \mathbb{R}^d and $f : O \rightarrow \mathbb{R} \cup \{-\infty\}$ be an upper semicontinuous function. Then there exists a unique maximal subharmonic function \hat{f} on O which is dominated by f . In other words, $\hat{f} \leq f$ and $g \leq \hat{f}$ for every $g \in \mathcal{SH}(O)$ with $g \leq f$. Furthermore, \hat{f} can be constructed as follows: For each $n \in \mathbb{N}$, define f_n on O by*

$$(3.3) \quad f_n(x) = \inf\{\mathbb{E}[f(F_n)] : (F_i)_{i=0}^n \text{ is a spherical martingale in } O \text{ with } F_0 = x \text{ and } F_n(\mathbb{P}) \text{ is compactly supported in } O\}.$$

Then, the sequence $(f_n)_{n=0}^\infty$ decreases pointwise to \hat{f} .

PROOF. We will use an equivalent form of (3.3):

$$f_0 = f \quad \text{and for } n \geq 1, \quad f_n(x) = \inf \left\{ \int f_{n-1}(x + ry) dU(y) \right\}$$

where the infimum is taken over all $r \geq 0$ such that $\{x + r\overline{B}\} \subset O$. Here, $\{x + r\overline{B}\}$ denotes the closed ball of radius r around x .

First note that the sequence $(f_n)_{n=0}^\infty$ is decreasing. To show the upper-semicontinuity of \hat{f} , we proceed inductively and assume f_{n-1} is upper semicontinuous. If $(x_k)_{k=0}^\infty$ in O is such that $\lim_{k \rightarrow \infty} x_k = x_0$, and $r \geq 0$ is such that $\{x_0 + r\overline{B}\} \subset O$, where \overline{B} is the closed unit ball, then there is k_0 such that $\{x_k + r\overline{B}\} \subset O$ for $k \geq k_0$. The upper semicontinuous function f_{n-1} is bounded above on the relatively compact set $\cup_{k=k_0}^\infty \{x_k + r\overline{B}\}$ and, for every $z \in \overline{B}$, $f_{n-1}(x_0 + rz) \geq \limsup_{k \rightarrow \infty} f_{n-1}(x_k + rz)$. Hence by Fatou's lemma,

$$\int f_{n-1}(x_0 + ry) dU(y) \geq \limsup_{k \rightarrow \infty} \int f_{n-1}(x_k + ry) dU(y) \geq \limsup_{k \rightarrow \infty} f_n(x_k).$$

Thus $f_n(x_0) \geq \limsup_{k \rightarrow \infty} f_n(x_k)$, hence showing that f_n and consequently \hat{f} is upper-semicontinuous.

Let now g be a subharmonic function on O with $g \leq f$. Again, inductively, assuming that $g \leq f_{n-1}$, then for $\{x_0 + r\overline{B}\} \subset O$,

$$\int f_{n-1}(x_0 + ry) dU(y) \geq \int g(x_0 + ry) dU(y) \geq g(x_0),$$

and so $f_n(x_0) \geq g(x_0)$. Hence $\hat{f} \geq g$. Finally, for $\{x_0 + r\overline{B}\} \subset O$, we get from monotone convergence

$$\hat{f}(x_0) = \lim_{n \rightarrow \infty} f_n(x_0) \leq \lim_{n \rightarrow \infty} \int f_{n-1}(x_0 + ry) dU(y) = \int \hat{f}(x_0 + ry) dU(y).$$

This shows that \hat{f} is subharmonic, and the proof of the lemma is complete. \square

Let now $\text{Lip}^*(O)$ be the space of all finite measures in O with finite first moments.

LEMMA 3.8. *Let $\mu, \nu \in \mathbb{P}_2(O)$ be such that $\mu \prec_{s_2(O)} \nu$, and consider the following subset of $\mathbb{P}_2(O)$,*

$$\Phi = \{F_n(\mathbb{P}) : (F_i)_{i=0}^n \text{ is a spherical martingale valued in } O \text{ with } F_0 \sim \mu\}.$$

Set $\tilde{\nu} := (1 + |x|)\nu$ and

$$\tilde{\Phi} = (1 + |x|)\Phi := \{\tilde{\sigma} \mid \tilde{\sigma} = (1 + |x|)\sigma \text{ for some } \sigma \in \Phi\}.$$

Then, $\tilde{\nu}$ is in the weak*-closure of $\tilde{\Phi}$ in $\text{Lip}^*(O)$.

PROOF. Observe first that Φ is convex. Indeed, if $(F'_i)_{i=0}^n$ and $(F''_i)_{i=0}^m$ are two spherical martingales with $F'_0 \sim \mu$, $F''_0 \sim \mu$, we may assume $n = m$, then define a spherical martingale $(F_i)_{i=0}^{n+1}$ by letting $F_0 = F_1 \sim \mu$ and for $1 \leq i \leq n$,

$$F_{i+1}(X_1, X_2, \dots, X_{i+1}) = \begin{cases} F'_i(X_2, \dots, X_{i+1}) & \text{if } X_1 \text{ is in the upper hemisphere,} \\ F''_i(X_2, \dots, X_{i+1}) & \text{if } X_1 \text{ is in the lower hemisphere.} \end{cases}$$

Clearly $F_{n+1}(\mathbb{P}) = \{F'_n(\mathbb{P}) + F''_n(\mathbb{P})\}/2$ and hence Φ is convex, and therefore $\tilde{\Phi} = (1 + |x|)\Phi$ is convex in $\text{Lip}^*(O)$.

If now the statement of the lemma were false, then by the Hahn-Banach theorem we can find a Lipschitz function f on O and real numbers $a < b$ such that

$$\int g d\nu \leq a, \text{ while } \int g \circ F_n d\mathbb{P} \geq b \text{ for every } F_n(\mathbb{P}) \in \Phi,$$

where $g(x) = (1 + |x|)f(x)$. But since every element in Φ has initial distribution μ , then by Lemma 3.7 and (3.3), we have $\int \hat{g} d\mu \geq b$ and therefore $a \geq \int g d\nu \geq \int \hat{g} d\nu \geq \int \hat{g} d\mu \geq b$, which is a contradiction. \square

Proof of Proposition 3.4. By Lemma 3.8, we have a sequence $\{\nu_n\} \subset \Phi$ such that $\int |x|^2 d\nu_n(x) \rightarrow \int |x|^2 d\nu(x)$. We know that $\nu_n = \text{Law}(B_{\tau_n})$ for a sequence of stopping times τ_n and a Brownian motion B with initial law μ . Hence in particular $\mathbb{E}\tau_n = \mathbb{E}|B_{\tau_n}|^2 = \int |x|^2 d\nu_n(x) \leq V$ for some constant V and for all n . The sequence (τ_n) is then tight, and it is standard that it has a convergent subsequence to a – possibly randomized – stopping time τ in such a way that $\mathbb{E}f(B_{\tau_k}) \rightarrow \mathbb{E}f(B_\tau)$ for every f continuous and bounded function on \mathbb{R}^d . (see for example [1] or [2]). In other words, $\text{Law}(B_{\tau_k}) \rightarrow \text{Law}(B_\tau)$, and therefore, $B_\tau \sim \nu$. Let τ_O be the first exit time of (B_t) from O and note that $\tau_k = \tau_k \wedge \tau_O$ by the definition of Φ , and therefore τ also satisfies $\tau = \tau \wedge \tau_O$. Notice also that $\mathbb{E}\tau = \mathbb{E}|B_\tau|^2 \leq V$ as well, hence the martingale $(B_{\tau \wedge t})_{t \geq 0}$ is bounded in L^2 . \square

Lastly, we prove the duality result announced in the introduction.

Proof of Theorem 1.6. By a standard argument, it is enough to prove the theorem for continuous cost c , so let us assume this. Clearly, the left-hand side is smaller than or equal to the right-hand side since for every $\pi \in \text{SMT}(\mu, \nu)$ (if exists) and for every $(\alpha, \beta, p) \in \mathcal{K}_c$, we have $\int p(x, y) d\pi(x, y) \geq 0$.

For the reverse inequality, we first note that Kantorovich duality for the standard optimal transport problem with a continuous cost $c(x, y) - p(x, y)$ yields

$$\sup_{(\alpha, \beta, p) \in \mathcal{K}_c} \left(\int \beta d\nu - \int \alpha d\mu \right) = \sup_{p \in \Psi} \inf_{\pi \in \Pi(\mu, \nu)} \int (c - p) d\pi.$$

Now it is standard to apply a min-max theorem (see e.g. [3, Theorem 2]) to interchange the order of inf and sup, and thereby obtain

$$(3.4) \quad \sup_{(\alpha, \beta, p) \in \mathcal{K}_c} \left(\int \beta d\nu - \int \alpha d\mu \right) = \inf_{\pi \in \Pi(\mu, \nu)} \sup_{p \in \Psi} \int (c - p) d\pi.$$

Note that the supremum over p can be finite only when $\int p(x, y) d\pi(x, y) \geq 0$, since otherwise we can replace p with λp for some $\lambda > 0$, making the value of the integral $\int (c - p) d\pi$ arbitrarily large. Therefore, by Theorem 1.5, the infimum in (3.4) can be restricted to $\pi \in \text{SMT}(\mu, \nu)$, and we obtain

$$\sup_{(\alpha, \beta, p) \in \mathcal{K}_c} \left(\int \beta d\nu - \int \alpha d\mu \right) = \inf_{\pi \in \text{SMT}(\mu, \nu)} \sup_{p \in \Psi} \int (c - p) d\pi.$$

Finally, since $0 \in \Psi$, the last expression is greater than or equal to

$$\inf \left\{ \int c(x, y) d\pi \mid \pi \in \text{SMT}(\mu, \nu) \right\}.$$

Together with Corollary 3.3, this completes the proof of the duality for the subharmonic martingale optimal transport problem. \square

4. Appendix: The radially symmetric monotonicity principle. In this section we use the dual attainment result [23] to provide a proof of Proposition 2.7.

Proof of Proposition 2.7. Let O be an open ball containing $\text{supp}(\mu + \nu)$. First, we can assume without loss of generality that $\text{supp} \mu \cap \text{supp} \nu = \emptyset$. Indeed, due to the assumption $\mu \wedge \nu = 0$ and $\mu(\partial \text{supp} \mu) = 0$, for each μ -a.e. x, x' , we can consider the restriction of μ to μ^+ on an open set containing x, x' outside $\text{supp} \nu$ and let ν^+ be the corresponding target measure under the stopping time τ . Then $\text{supp} \mu^+ \cap \text{supp} \nu^+ = \emptyset$. It will be sufficient to prove the desired conclusion of the proposition for those $x, x' \in \text{supp} \mu^+$, $\text{supp} \varphi_y \subset \text{supp} \nu^+$, and $y' \in \text{supp} \nu^+$.

Now, from [23, Theorem 4.1 and the proof of Theorem 7.1 (1)] we have the existence of the optimizers for the dual problem. Here, note that the theorems in [23] can be used because $\text{supp} \mu \cap \text{supp} \nu = \emptyset$, the function $c(x, y) = \pm|x - y|^\alpha$, $\alpha > 0$, is C^2 for those $x, y \in \text{supp} \mu \times \text{supp} \nu$; this cost can be modified by adding an appropriate subharmonic function $h(y)$, with large Δh , to an equivalent cost $y \mapsto c(x, y) + h(y)$ with $0 \leq \Delta_y[c(x, y) + h(y)]$ which is C^2 on $\text{supp} \mu \times \text{supp} \nu$. Then, as in the proof of [23, Theorem 7.1 (1)] this can be extended to a C^2 subharmonic function on $O \times O$ without changing the optimal solution. The result [23, Theorem 4.1] can be applied to this cost to get the dual optimizers, first $\beta \in H_0^1(O) \cap LSC(O)$, then

$$(4.1) \quad J(x, y) := \sup_{\xi \leq \tau_O} [\beta(B_\xi^y) - c(x, B_\xi^y)],$$

where ξ is a stopping time for Brownian motion, and τ_O is the exit time from the domain O . Set $\alpha(x) = J(x, x)$ and $p(x, y) = J(x, x) - J(x, y)$.

From the radial symmetry of μ and ν , and the expression of the dual problem, we can assume that the dual optimizer β is radially symmetric. To see this, recall that \mathfrak{M} is the group of $d \times d$ orthogonal matrices and \mathcal{H} is the Haar measure on \mathfrak{M} . Then, for each $M \in \mathfrak{M}$,

$$J(Mx, My) = \alpha(Mx) - p(Mx, My) = \sup_{M\xi \leq \tau_O} [\beta(B_{M\xi}^{My}) - c(Mx, B_{M\xi}^{My})].$$

Since both Brownian motion and the cost $c(x, y) = \pm|x - y|^\alpha$ are invariant under the orthogonal group (isometries of \mathbb{R}^d) we see that for

$$\bar{J}(x, y) = \int_{M \in \mathfrak{M}} J(Mx, My) d\mathcal{H}(M) \quad \text{and} \quad \bar{\beta}(y) = \int_{M \in \mathfrak{M}} \beta(My) d\mathcal{H}(M),$$

it holds that

$$\bar{J}(x, y) = \sup_{\xi \leq \tau_O} [\bar{\beta}(B_\xi^y) - c(x, B_\xi^y)].$$

Notice that $\bar{\beta}$, $\bar{\alpha}(x) := \bar{J}(x, x)$ and $\bar{p}(x, y) := \bar{J}(x, x) - \bar{J}(x, y)$ are admissible and they have the same optimal dual value as $\int \beta d\nu - \int \alpha d\mu$, thus they are dual optimizers.

The rest of the proof is similar to that of [23, Theorem B.1]. We will use the radial symmetry of β . We notice that the desired conclusion of our theorem is equivalent to the following:

CLAIM 4.1. *Let π be the measure with $(B_0, B_\tau) \sim \pi$. Consider a randomized stopping time ξ with $B_0 \sim \mu$, $(B_0, B_\xi) \sim \sigma$ where the disintegration $\sigma(dx, dy) = \sigma_x(dy)\mu(dx)$ satisfies*

$$(4.2) \quad 0 \prec_{s(O)} \sigma_x \prec_{s(O)} \pi_x \text{ for each } \mu\text{-a.e. } x,$$

which means that $\sigma \leq \tau$, and $\tau - \sigma$ is a (randomized) stopping time.

Then, for π -a.e. (x', y') and σ -a.e. (x, y) , it holds that if $|y| = |y'|$, then the stopping time $\tau - \xi$ restricted to paths with $B_\xi = y$ satisfies

$$c(x', y') + \mathbb{E}[c(x, B_{\tau-\xi}^y)] \leq \mathbb{E}[c(x', B_{\tau'}^{y'})] + c(x, y),$$

for any stopping time τ' such that $B_{\tau'}^{y'} \sim \varphi_{y'} \cong_R \psi_y \sim B_{\tau-\xi}^y$.

To prove this claim notice that from [23, Theorem 4.7], we have for π^* -a.e. (x', y') and σ -a.e. (x, y) ,

$$(4.3) \quad J(x', y') = \beta(y') - c(x', y'),$$

$$(4.4) \quad J(x, B_{\tau-\xi}^y) = J(x, B_\tau^x) = \beta(B_\tau^x) - c(x, B_\tau^x) = \beta(B_{\tau-\xi}^y) - c(x, B_{\tau-\xi}^y).$$

On the other hand, from the definition (4.1), we have

$$(4.5) \quad J(x', y') \geq \mathbb{E}[\beta(B_{\tau'}^{y'}) - c(x', B_{\tau'}^{y'})] = \mathbb{E}[\beta(B_{\tau-\xi}^y) - c(x', B_{\tau'}^{y'})]$$

where the first inequality holds for any τ' and the second one holds in particular for those with $B_{\tau'}^{y'} \sim \varphi_{y'} \cong_R \psi_y \sim B_{\tau-\xi}^y$ due to radial symmetry of β . Notice that from [23, Theorem 4.7 (4.4)] we have

$$(4.6) \quad \mathbb{E}[J(x, B_{\tau-\xi}^y)] = J(x, y).$$

Moreover,

$$J(x, y) \geq \beta(y) - c(x, y) = \beta(y') - c(x, y)$$

where the last equality uses the radial symmetry of β and the fact that $|y| = |y'|$.

Taking the expectation in (4.4), we see that the left hand sides of (4.3), (4.4) are equal to the left hand sides of inequalities (4.5) and (4.6). Now, subtract the sum of the two equations from the sum of the two inequalities and cancel the terms $\beta(y)$ and $\mathbb{E}[\beta(B_{\tau-\xi}^y)]$, to obtain

$$0 \geq -c(x, y) - \mathbb{E}[c(x', B_{\tau'}^{y'})] + c(x', y') + \mathbb{E}[c(x, B_{\tau-\xi}^y)],$$

hence completing the proof. \square

REFERENCES

- [1] J. R. Baxter and R. V. Chacon, Potentials of stopped distributions. *Illinois J. Math.*, 18:649–656, (1974).
- [2] M. Beiglböck, A.M.G. Cox, and M. Huesmann, Optimal transport and Skorokhod embedding. *Inventiones mathematicae*, Volume 208, Issue 2 (2017) 327–400.
- [3] M. Beiglböck, P. Henry-Labordère, and F. Penkner, Model-independent bounds for option prices—a mass transport approach. *Finance and Stochastics*, Volume 17 (2013) 477–501.
- [4] M. Beiglböck, P. Henry-Labordère, and N. Touzi, Monotone martingale transport plans and Skorohod Embedding. *Stochastic Processes and their Applications*, Volume 127, Issue 9 (2017) 3005–3013.
- [5] M. Beiglböck and N. Juillet, On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.*, Volume 44 (2016) 42–106.
- [6] M. Beiglböck, M. Nutz, and N. Touzi, Complete duality for martingale optimal transport on the line. *Ann. Probab.*, Volume 45, Number 5 (2017) 3038–3074.
- [7] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions. *Communications on pure and applied mathematics*, Volume 44, Issue 4 (1991) 375–417.
- [8] S. Bu and W. Schachermayer, Approximation of Jensen measures by image measures under holomorphic functions and applications. *Trans. Amer. Math. Soc.*, Volume 331 (1992) 585–608.
- [9] R. V. Chacon and J. B. Walsh, One-dimensional potential embedding. In *Séminaire de Probabilités, X. Lecture Notes in Math.*, Vol. 511. Springer, Berlin (1976), 19–23. MR445598.
- [10] G. Choquet, Forme abstraite du théorème de capacabilité. *Ann. Inst. Fourier. Grenoble*, Volume 9 (1959) 83–89.
- [11] A.M.G. Cox, J. Oblój, and N. Touzi, The Root solution to the multi-marginal embedding problem: an optimal stopping and time-reversal approach. *Probability Theory and Related Fields* (Open Access), (2018).
- [12] H. De March and N. Touzi, Irreducible convex paving for decomposition of multi-dimensional martingale transport plans. <https://arxiv.org/abs/1702.08298> (2017).
- [13] Y. Dolinsky and H.M. Soner, Martingale optimal transport and robust hedging in continuous time. *Probab. Theory Relat. Fields*, 160 (2014) 391–427.
- [14] Y. Dolinsky and H.M. Soner, Martingale optimal transport in the Skorokhod space. *Stochastic Processes and their Applications*, 125(10) (2015) 3893–3931.
- [15] G.A. Edgar, Complex martingale convergence. *Banach Spaces, Volume 1166 of the series Lecture Notes in Mathematics*, 38–59.
- [16] G.A. Edgar, Analytic martingale convergence. *Journal of Functional Analysis*, Volume 69, Issue 2, November 1986, 268–280.
- [17] N. Falkner, On Skorokhod embedding in n-dimensional Brownian motion by means of natural stopping times. In *Séminaire de Probabilités, XIV, volume 784 of Lecture Notes in Math.*, 357–391. Springer, Berlin, (1980). MR580142.
- [18] G.B. Folland. *Real analysis: modern techniques and their applications*. John Wiley & Sons, 2013.
- [19] A. Galichon, P. Henry-Labordere, and N. Touzi, A Stochastic Control Approach to No-Arbitrage Bounds Given Marginals, with an Application to Lookback Options. *Ann. Appl. Probab.*, Volume 24, Number 1 (2014) 312–336.
- [20] W. Gangbo and R.J. McCann, The geometry of optimal transportation. *Acta Math*, Volume 177, Issue 2 (1996) 113–161.
- [21] N. Ghoussoub, Y-H. Kim, and T. Lim, Structure of optimal martingale transport in general dimensions. *Annals of Probability*, Vol. 47, No. 1, (2019) 109–164

- [22] N. Ghoussoub, Y-H. Kim, and T. Lim, Optimal Skorokhod Embedding for radially symmetric marginals arXiv:1711.02784
- [23] N. Ghoussoub, Y.-H. Kim, and A.Z. Palmer. A solution to the Monge transport problem for Brownian martingales. arXiv:1903.00527
- [24] N. Ghoussoub, Y.-H. Kim, and A.Z. Palmer, PDE Methods for Optimal Skorokhod Embeddings, *Calculus of Variations and PDE*, 58: 113.: 1-33.
- [25] N. Ghoussoub and B. Maurey, Plurisubharmonic martingales and barriers in complex quasi-Banach spaces. *Annales de l'institut Fourier*, Volume 39, Issue 4 (1989) 1007–1060.
- [26] G. Guo, X. Tan, and N. Touzi, Optimal Skorokhod embedding under finitely-many marginal constraints. *SIAM J. Control Optim.*, 54(4), (2016) 2174–2201.
- [27] D. Hobson, Robust hedging of the lookback option, *Finance and Stochastics*, 2 (1998) 329–347.
- [28] D. Hobson, The Skorokhod embedding problem and model-independent bounds for option prices. In *Paris-Princeton Lectures on Mathematical Finance 2010*, volume 2003 of *Lecture Notes in Math*, Springer, Berlin (2011) 267–318.
- [29] D. Hobson and M. Klimmek, Robust price bounds for the forward starting straddle. *Finance and Stochastics*, Volume 19, Issue 1 (2014) 189–214.
- [30] D. Hobson and A. Neuberger, Robust bounds for forward start options. *Mathematical Finance*, Volume 22, Issue 1 (2012) 31–56.
- [31] J. A. Johnson, Banach spaces of Lipschitz functions and vector-valued Lipschitz functions. *Bull. Amer. Math. Soc.*, Volume 75, Number 6 (1969), 1334–1338.
- [32] S. Källblad, X. Tan and N. Touzi, Optimal Skorokhod embedding given full marginals and Azema-Yor peacocks. *Ann. Appl. Probab.*, Volume 27, Number 2 (2017), 686–719.
- [33] F. Ledrappier and L.-S. Young, The Metric Entropy of Diffeomorphisms: Part I: Characterization of Measures Satisfying Pesin's Entropy Formula. *Annals of Mathematics*, Second Series, Vol. 122, No. 3 (1985), 509–539.
- [34] T. Lim, Optimal martingale transport between radially symmetric marginals in general dimensions. <https://arxiv.org/abs/1412.3530>, (2016).
- [35] I. Monroe, On embedding right continuous martingales in Brownian motion. *The Annals of Mathematical Statistics*, Vol. 43, No. 4, 1293–1311 (1972).
- [36] J. Oblój, The Skorokhod embedding problem and its offspring *Probability Surveys*, 1 (2004) 321–392.
- [37] J. Oblój and P. Siorpaes, Structure of martingale transports in finite dimensions. <https://arxiv.org/abs/1702.08433> (2017).
- [38] J. Oblój and P. Spoida, An Iterated Azéma-Yor Type Embedding for Finitely Many Marginals. *Ann. Probab.*, to appear. (2016)
- [39] H. Rost, The Stopping Distributions of a Markov Process. *Inventiones mathematicae*, 14, 1–16 (1971).
- [40] A.V. Skorokhod, Studies in the Theory of Random Processes (Translated from the Russian by Scripta Technica Inc.). *Addison-Wesley Publishing Co. Inc.*, Reading (1965)
- [41] V. Strassen, The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36 (1965) 423–439.
- [42] C. Villani, *Topics in optimal transportation*, Volume 58, *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
- [43] J.L. Walsh, The approximation of harmonic functions by harmonic polynomials and by harmonic rational functions. *Bull. Amer. Math. Soc.* (2) 35 (1929) 499–544.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, V6T 1Z2 CANADA
E-MAIL: nassif@math.ubc.ca
yhkim@math.ubc.ca,

INSTITUTE OF MATHEMATICAL SCIENCES
SHANGHAI TECH UNIVERSITY
No. 393, HUAXIA MIDDLE ROAD, PUDONG NEW AREA, SHANGHAI 201210
E-MAIL: TLIM@shanghaitech.edu.cn